

## Image Recognition Apparatus, Method and Program Product

[0001] This application is based on Japanese Patent Application No. 2000-136301 filed on May 9, 2000, the contents of which are hereby incorporated by reference.

### BACKGROUND OF THE INVENTION

#### 1. Field of the Invention

[0002] The present invention relates to an image recognition apparatus, method and program product that have a capability of maintaining the original layout when character code data, which are obtained by recognizing character codes from character images included in input image data, are disposed using a particular font.

#### 2. Description of the Related Art

[0003] The image recognition apparatus is to analyze image data inputted by scanning an original document by a scanner and the like, extracting each region that contains character images, photograph images or diagram images by each attribute, and conducting recognition process on each region according to its attribute.

[0004] In recent years, image recognition apparatuses, having a capability not only to output character code data obtained by recognizing character codes from character images

but also to output them maintaining the original document's layout, have been proposed.

[0005] For example, an image recognition apparatus is known, which is capable of recognizing from character images font sizes, font types, styles (bold, italics, etc.) and maintaining the original document's layout by outputting character code data including these attributes of characters.

[0006] Another image recognition apparatus is also known, which is capable of embedding photographic and diagram images as bit map data into output image data, or recognizing page formats to make it possible to describe them in terms of layout frame styles, so that various styles of files can be prepared.

[0007] However, those image recognition apparatuses can recognize only a limited level of font types as character attributes from character images. Moreover, since numerous types of fonts are used on actual documents, it is not realistic to recognize all fonts. Moreover, even if all fonts are accurately recognized and the recognized font is specified and outputted, it is not possible to receive the data with accurate fonts unless the data receiving side is prepared with such fonts.

[0008] Therefore, numerous types of fonts existing on original documents (hereinafter "input fonts") have no choice but to be reduced to a limited number of general fonts that

are used for outputting character code data (hereinafter "output fonts").

[0009] Consequently, it is often impossible to output character code data to have the same character layout as in the original document considering the fact that the width and height of a character vary with the type of a font used when the character code obtained from a character image is outputted in a specific output font.

[0010] It is therefore difficult to deal with a large varieties of documents even if those image recognition apparatuses mentioned above are equipped with capabilities to maintain document layouts. More specifically, there are problems as described below depending on the fonts and layouts of texts and graphics on an original document, when character code data obtained by recognizing them from character images are outputted in specific output fonts.

[0011] For example, if the character width of an output font is greater than the character width of an input font, the character image outputted based on the character code data in the output font may not fit within the character region in the image data obtained by scanning the document, where the character image exists. This causes a problem such that characters and photographs overlap each other, or characters overlap each other on the output image. In some cases, the

number of text lines increases in the process, so that the result of scanning one page may result in two pages of output image. On the other hand, when a process of separating a text from a photograph is performed, the photograph may get outputted on a separate page.

#### SUMMARY OF THE INVENTION

**[0012]** It is an object of the present invention to provide an improved image recognition apparatus, method and program to solve the problems described above.

**[0013]** It is a further object of the present invention to provide an image recognition apparatus, method and program that allow to layout output character images maintaining the original layout and without overlapping other character images or graphic images within one page even when character code data, which are obtained by recognizing character codes from character images included within input image data, are disposed according to a certain specific font.

**[0014]** According to an aspect of the invention we provide an image recognition apparatus comprising: an input unit for inputting image data; and a processor for executing a process comprising the steps of: 1) detecting a character region where character images exist from image data input via said input unit; 2) recognizing character images in said character region

to obtain character code data; 3) converting said character code data into output character images; 4) judging whether said output character images fit within said detected character region when said output character images are outputted; and 5) enlarging said character region within a specified range when said output character images do not fit into said character region and judging whether said output character images fit within the enlarged region.

**[0015]** According to another aspect of the invention we provide an image recognition method comprising the steps of: 1) inputting image data; 2) detecting a character region where character images exist from the input image data; 3) recognizing character images in said character region to obtain character code data; 4) converting said character code data into output character images; 5) judging whether said output character images fit within said detected character region when said output character images are outputted; and 6) enlarging said character region within a specified range when said output character images do not fit into said character region and judging whether said output character images fit within the enlarged region.

**[0016]** According to still another aspect of the invention we provide a computer readable program product for recognizing images, said program product causing the computer to execute



process of the first embodiment;

[0020] Fig. 3 is a flow chart showing output adjustment process shown in Fig. 2;

[0021] Fig. 4 is a diagram showing an example document;

[0022] Fig. 5 is a diagram showing the result of detecting character regions by means of conducting layout analysis on image data obtained by scanning the document shown in Fig. 4;

[0023] Fig. 6 is a diagram for describing a method of judging whether a line in a character region is a continuous line;

[0024] Fig. 7 is a diagram for describing a method of judging whether a prescribed amount of unused area exists;

[0025] Fig. 8 is an outline block diagram showing the constitution of an information equipment system containing the second embodiment of the image recognition apparatus according to the present invention;

[0026] Fig. 9 is a flow chart showing the image recognition process of the second embodiment;

[0027] Fig. 10 is a flow chart showing the output adjustment process shown in Fig. 9;

[0028] Fig. 11 is a diagram showing an example of detecting character regions and character background regions by means of conducting layout analysis on image data obtained by scanning a document; and

[0029] Fig. 12 is a diagram for describing a method of judging where each line is left aligned, right aligned or centered in a character background region.

#### DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

[0030] The embodiment of this invention will be described below with reference to the accompanying drawings.

[0031] Fig. 1 is an outline block diagram showing the constitution of an information equipment system containing a digital copying machine 10 as the first embodiment of the image recognition apparatus according to the present invention.

[0032] As shown in Fig. 1, an information equipment system 100 comprises the digital copying machine 10, a computer 50 and other information equipment interconnected on a network N.

[0033] The digital copying machine 10 having a function as an image recognition apparatus comprises a main control unit 11 that integrally controls various parts of the digital copying machine, a scanner 12 that scans documents to obtain image data, an image processing unit 13 that analyzes the image data to obtain character code data by recognizing character codes from character images and conducts processes of data compression, etc., a print unit 14 that prints specific



data on paper, a ROM 15 that stores specific programs and data prepared in advance, a RAM 16 that stores data temporarily, an operating panel 17 for displaying various information or by which the user issues instructions, and an interface (I/F) 18 for transmitting or receiving commands and data between the computer 50 and other information equipment.

[0034] The digital copying machine 10 and the computer 50 are capable of exchange data via the network N. For example, the digital copying machine 10 can transmit the image data which is obtained by scanning a document by the scanner 12 and is image processed by the image processing unit 13 to the computer 50. It is also possible for the digital copying machine 10 to receive image data from the computer 50 and print it on paper by the printer unit 14.

[0035] The digital copying machine 10 has a printing function that uses the printer unit 14 independently and a scanner function which uses the scanner unit 12 independently in addition to the copying function wherein image data obtained by scanning a document by the scanner 12 is image processed depending on the necessity and printed in the printer unit 14.

[0036] The image data obtained by scanning a document by the scanner 12 is, for example, sent to the image processing unit 13, where the image processing unit 13 recognizes

character codes from character images to convert them into character code data, and prepares files of a specified format. The prepared file is transmitted to the computer 50, if there is an instruction to take it into the computer 50 from the operating panel 17 of the digital copying machine 10, and transmitted to the printing unit 14 if there is an instruction to make copy from the operating panel 17. The prepared file can also be transmitted attached to electronic mail via the network N.

**[0037]** As shown in Fig. 1, the image processing unit 13 comprises a layout analysis unit 31, a character recognition unit 32, a font recognition unit 33, a character region change unit 34, a character size change unit 35, and a file generating unit 36.

**[0038]** The layout analysis unit 31 analyzes the input image data, extracts the character region where character images exist, the photograph region where graphic images, i.e., non-character images, exist, and the diagram region where a mixture image of character images and graphic images, such as graphs and charts, exist, respectively depending on the attributes of the image data.

**[0039]** Character images in character regions are analyzed by the character recognition unit 32, which will detect character codes and convert them into character code data.

Graphic images within the photographic region are stored, for example, on the file in bit map data as received as the input image data. As to character images in a diagram region, the user determines whether they are to be converted into the character code data or not by selecting the setting on the panel 17.

[0040] The extraction of character region, photograph region and diagram region is conducted by the known method. For example, the extraction of the character region is done by measuring the sizes of the areas that are graphically connected in the image data and using a criterion that the sizes of character images are generally smaller than those of other graphic images. Although the layout analysis in this embodiment is done by multi-value image data, it can be done by using the binary image data.

[0041] The character recognition unit 32 obtains character code data by recognizing a character code from each character image, and it also recognizes the position information of each character image. The position information of each character image is obtained as, for example, the coordinate of the top left and the bottom right corner of a rectangular shape circumscribed about each character image. The recognition of the character code from the character image is, for example, done based on the degree of matching between

the characteristic value of each character image and the dictionary patterns stored in the ROM 15.

[0042] The font recognition unit 33 recognizes the type of font and the font size from the character image. For example, the font is identified by classifying it into a specified typical several output fonts using the same as the method of recognition of the character code data as described above.

[0043] The character region changing unit 34 can adjust the size of a character region, to which the character image outputted based on the character code data is disposed, to expand from a character region in the image data obtained by scanning the document. The character size change unit 35 can also reduce the size of a character image outputted based on the character code data.

[0044] The file generating unit 36 generates a specific format of file by applying a specific process to the input image data such as recognizing a character code from the character image and converting it to a character code data.

[0045] Next, the operation of a digital copying machine having a function as the image recognition apparatus constituted as above.

[0046] Fig. 2 is a flow chart showing the image recognition process and Fig. 3 is a flow chart of the output adjustment process shown in Fig. 2.



**[0051]** Next, character code data are obtained as character codes are recognized by the character recognition unit 32 from individual character images. Also, position information of each character image is recognized (S14). Further, the type of font and font size are recognized from the character image by the font recognition unit 33 (S15). The font is recognized from typical several output fonts prepared in the image recognition apparatus.

**[0052]** When the character code data thus obtained are arranged in the output image in the output font recognized by the font recognition unit 33, it makes a judgment whether the character image sticks out of a character region in the image data inputted by scanning the document, and the output adjustment process is done according to the judgment result (S16). Although a font recognized by the font recognition unit 33 is normally applied as the output font, it is also possible to constitute in such a way that the user can arbitrarily set up.

**[0053]** In this output adjustment process, the pitch of the lines within the character region is calculated as shown in Fig. 3 (S21). The line pitch is calculated by dividing the height of the character region with the number of lines. The height and width of the character region can be obtained from the layout analysis (S12). Moreover, the number of lines can

be obtained by checking the x-coordinate (lateral coordinate) of the circumscribed rectangular shape as position information of each character image one character after another and detecting a line feed when the x-coordinate makes a change in the negative direction (opposite direction). However, this method does not allow us to calculate the line pitch if the particular character region consists of only one line. In such a case, a decision is made not to set up the line pitch.

**[0054]** Next, a judgment is made whether the lines in the character region are continuous lines (S22). The character region A of Fig. 5 shows a case where the lines are continuous and the character region B of the same shows a case where the lines are independent lines.

**[0055]** Fig. 6 shows a method for identifying whether a line in a character region is a continuous line. The judgment whether the lines in the character region is continuous can be made by detecting the variations of the starting point and the ending point of each line. More specifically, if the averages of the variations of the starting point and the ending point of each line of a character image that consists of multiple lines in relation to the adjacent lines are greater than a specified value, it is judged that the lines in the character region are independent, and that the lines are

continuous if they are smaller than the specified value. With reference to Fig. 6, let  $S_i$  ( $i = 1, 2, \dots$ ) be the x-axis coordinate of the starting point of the  $i$ -th line,  $E_i$  ( $i = 1, 2, \dots$ ) be the x-axis coordinate of the ending point of the  $i$ -th line, and  $N$  be the number of lines, and calculate:

$$\Sigma \{ (S_{i+1} - S_i) + (E_{i+1} - E_i) \} / N$$

and determine whether the lines within the particular character region are continuous lines based on the size of the calculated value. However, the calculation is performed on the assumption that a line  $L$  has no variation of the ending position if the line includes a character image that indicates the end of the sentence such as a period as shown in Fig. 6.

[0056] The above steps will provide information on the height and width of the character region, line pitch and number of lines, judgment result on whether the lines in the character region are continuous, character code, font size, type of font, position of each character image (coordinate of the circumscribed rectangle), and font metric data used for outputting. The font metric data is the dimensional data concerning the character width, the character height, etc., of a font. The character width is used as the font metric data in this embodiment.

[0057] If the lines in a character region are continuous



lines, it is acceptable as long as the output character image based on the character code data obtained by recognizing the character codes from character image can be placed as a whole within the character region of the image data inputted by scanning a document, even if the character widths are different because the output font recognized by the font recognition unit 33 is different from the input font of the characters on the document. On the other hand, if the lines within the character region are not continuous lines but are independent lines such as bulleted items, it would be inappropriate to add a part of a line to another line, as it would make it difficult for the reader to understand the sentence. Therefore, the layout of the output character images which are based on the character code data obtained by recognizing them from character images varies depending on whether the lines within the character region are continuous lines (YES at S23) or the lines are independent lines (NO at S23).

**[0058]** When the lines in a character region are continuous lines (YES at S23), the width of the output character image lines that can be arranged within the character region is calculated (S24).

**[0059]** Let  $X(i)$  be the width of the character region in question ( $i$ -th character region of multiple character regions where  $i = 1, 2, \dots$ ) and  $N(i)$  be the number of lines in the

same region, the width of the output character image lines  $Lorg(i)$  in the particular character region can be calculated as follows:

$$Lorg(i) = X(i) \cdot N(i)$$

[0060] On the other hand, the width,  $Lout(i)$ , of the output character image lines that are actually disposed can be calculated as follows, assuming  $Wout(i)$  be the character width of the font available from the font metric data used for the output and  $Nc(i)$  be the number of the character code data:

$$Lout(i) = Wout(i) \cdot Nc(i)$$

[0061] If a font, in which the character width varies with each character code data, is to be used, it is necessary to add the character width  $Wout(ci)$  corresponding to each character code data, so that:

$$Lout(i) = \sum Wout(i)$$

[0062] Next, a judgment is made if the output character image sticks out of the character region of the image data inputted by scanning the document, when the character code data obtained in the above process is outputted using the output font as the character image (S25).

[0063] If  $Lout(i) \leq Lorg(i)$ , it is judged that the output character image does not stick out of the character region of the input image data (YES at S25), the character code data

will be outputted in the output font as the character image (S31).

[0064] On the other hand, if  $L_{out}(i) > L_{org}(i)$ , it is judged that the output character image does stick out of the character region of the input image data (NO at S25), and a judgment is made whether the character region in which the output character image is disposed can be expanded from the character region in the image data inputted by scanning the document. In other words, a judgment is made whether there is a specified amount of unused portion corresponding to the base region of the document that does not contain either character images or graphic images, i.e., non-character images, adjacent to the character region in the input image data (S26).

[0065] Fig. 7 is a diagram for describing a method of judging whether a specified amount of unused area exists. A judgment is made whether an unused area exists expanding the area of check with a fixed width (length) increment in four directions, i.e., left, right, above and below, adjacent to a certain character region A in the input image data.

[0066] For example, a region M1 adjacent to the character region A on the right side is checked first. The width of the region M1 is chosen to be approximately equal to the character width  $W_{out}$  of the output font. Here, the density

of each pixel of the region M1 is checked using the multi-value image data, and it is judged that the region M1 is an unused area that does not contain either character images or graphic images, i.e., non-character images. It is also possible to use binary image data for making a judgment whether the region M1 is unused area.

**[0067]** If the region M1 is determined to be an unused area, another judgment is made whether the output character image sticks out of the character region when the character region is expanded to a region that includes said unused area. If the region M1 is not an unused area, another expansion is not made to the right of the character region A. Similarly, judgments are made sequentially for the regions M2 through M8. The width of the regions adjacent above and below of the character region A is chosen to be approximately equal to the height of the output font.

**[0068]** It is judged that a specified amount of unused area exists if the output character image fits into the expanded character region when the character region is sequentially expanded into the region that contains the unused area in this manner (YES at S26). The character region is expanded to a new region that includes said unused area by means of the character region change unit 34, and the character code data is outputted in the output font as the character image

(S27) .

**[0069]** In the meantime, if no unused area of a specified amount existed (NO at S26), the output character image based on the character code data within said character region is disposed by reducing the character size by means of the character size change unit 35 so that the output character image can fit within the character region (S30). In order to reduce the character size, a method of reducing the font size is used in this embodiment. However, it is also possible to use a method of specifying a smaller scaling factor for the entire character region.

**[0070]** If a line within the character region is an independent line (NO at S23), the same processing as described above for continuous lines is applied line by line. In other words, the width of the output character image lines is calculated to make it fit within a line region obtained by dividing the character region into the number of lines (S29). A judgment is made whether an unused area exists, expanding the area of check with a fixed width (length) increment in two directions, i.e., left and right, adjacent to said line region. Since other points are the same as in the case of continuous lines, detailed descriptions are not repeated here.

**[0071]** If there are multiple character regions exist in

the input image data, said output adjustment processes (S16, S21-S31) are performed on all character regions (S28).

[0072] When the output adjustment process is completed, the system returns to the main flow chart and the character code data in the output adjustment processed character regions are outputted in a specified format of file together with graphic images, etc., for example, in other photographic regions (S17). The output positions of the character image, etc., are specified depending on the file format.

[0073] As shown above, it is possible to dispose the output character images within the same page avoiding overlaps with other output character images and/or graphic images while maintaining the original layout even in the case of outputting the character code data, which are obtained by recognizing them from the character images in the input image data, using a specific font.

[0074] Fig. 8 is an outline block diagram showing the constitution of an information equipment system containing a digital copying machine as the second embodiment of the image recognition apparatus according to the present invention. Differences from the first embodiment will be mainly described below.

[0075] In this embodiment, as shown in Fig. 8, an image processing unit 13a comprises a layout analysis unit 31, a

character recognition unit 32, a font recognition unit 33, a character region change unit 34, a character size change unit 35, and file generating unit 36, as well as a character background region detection unit 37, a character position detection unit 38 and a layout style specifying unit 39.

[0076] Fig. 9 is a flow chart showing the image recognition process of the second embodiment. Fig. 10A and Fig. 10B show a flow chart showing the output adjustment process shown in Fig. 9.

[0077] After the OCR pre-processing (S41) and the layout analysis (S42) are performed similar to those in the first embodiment (S11 and S12 of Fig. 2), the character background region detection will be performed (S43).

[0078] Fig. 11 is a diagram showing an example of detecting character regions and character background regions by means of conducting layout analysis on image data obtained by scanning a document.

[0079] As shown in Fig. 11, the layout analysis unit 31 detects and extracts, for example, rectangular character regions F, G, etc., where character images exist. The character background region detection unit 37 detects character background regions H and I that contain the character regions F and G respectively and serve as their backgrounds, while existing inside another different background region





**[0081]** In the output adjustment process, as shown in Fig. 10A and Fig. 10B, the line pitch calculation (S51) and a judgment whether lines in a character region are continuous lines (S52) will be performed similar to the first embodiment (S21 and S22 of Fig. 3), which will then be followed by a judgment whether character regions are included in respective character background regions (S53).

**[0082]** For example, if a character region is not included in a character background region as the character region K shown in Fig. 11 (NO at S53), similar processes as the first embodiment (S23-S27, S29-S31 of Fig. 3) will be performed (S65).

**[0083]** If a character region is included in a character background region (YES at S53) and lines in the character region are independent lines, the layout of each line within the character background regions will be detected (S54). In other words, the character position detection unit 38 detects whether the particular lines are left aligned, right aligned or center aligned.

**[0084]** Fig. 12 is a diagram for describing a method of judging where each line is left aligned, right aligned or centered in a character background region. This judgment is conducted by calculating the sum of the deviations of the x-axis coordinates of the starting point, ending point and

center point of each line from their averages within the particular character region, and to see which one has the least total of the deviations.

**[0085]** Let  $L_j$  ( $j = 1, 2, \dots$ ) be the start point x-axis coordinate of the  $j$ -th line,  $R_j$  ( $j = 1, 2, \dots$ ) be the end point x-axis coordinate of the  $j$ -th line,  $C_j$  ( $j = 1, 2, \dots$ ) be the center point x-axis coordinate of the  $j$ -th line, and  $L$ ,  $R$  and  $C$  be the averages, the layout of the character image line within the particular character background region can be judged as being:

left aligned if  $\text{Min} \{ \sum (L_j - L), \sum (R_j - R), \sum (C_j - C) \} = \sum (L_j - L);$

right aligned if  $\text{Min} \{ \sum (L_j - L), \sum (R_j - R), \sum (C_j - C) \} = \sum (R_j - R);$  and

centered if  $\text{Min} \{ \sum (L_j - L), \sum (R_j - R), \sum (C_j - C) \} = \sum (C_j - C).$

**[0086]** If the lines within the character region are continuous (YES at S55), the width of the output character image lines that can be arranged within the character region is calculated (S56) same as in the first embodiment (S24 of Fig. 3). On the other hand, if the lines within a character region are independent (NO at S55), the width of the output character image line is calculated to make it fit within a line region obtained by dividing the character region into

the number of lines (S62) same as in the first embodiment (S29 of Fig. 3).

**[0087]** Then, when the thus obtained character code data are outputted in the output font used for outputting as the character image, a judgment is made whether the outputted character image sticks out of the character region in the image data inputted by scanning the document (S57). If the outputted character image does not stick out of the character region in the inputted image data (YES at S57), the character code data are outputted (S64) as the character image using the output font into said character region same as in the first embodiment (S31 of Fig. 3).

**[0088]** In the second embodiment, the process to be taken when the output character image sticks out of the character region in the inputted image data (NO at S57) is different from the first embodiment. In other words, it is not to judge whether there is any unused area adjacent in the left, right, top and bottom directions, but rather it is to judge whether an output character image fits within an expanded character region when the character region where the output character image is to be disposed is expanded in the right and left or top and bottom directions within the range of the character background region (S58).

**[0089]** If an output character image fits within a character

background region (NO at S58), the character region where the output character image is placed is expanded within the range of the character background region by means of the character region change unit 34, so that the character code data are outputted as a character image using an output font (S59).

**[0090]** On the other hand, even in a case where a character region where an output character image is placed is expanded until it matches with the corresponding character background region, if the output character image does not fit inside the character background region (YES at S58), the output character image within said character region will be disposed by reducing the character size by means of the character size change unit 35 so that the entire output character image fits within the character region (S63).

**[0091]** Next, when the lines within a character region are independent lines, the layout style for each line is specified by the layout style specifying unit 39 based on the detection result of the character position detection unit 38 (S60).

**[0092]** If multiple character regions exist within input image data, the above-mentioned output adjustment processes (S47, S51-S65) will be performed on all the character regions (S61).

**[0093]** When the output adjustment processes are finished,

the system returns to the main flow chart, the character code data within the character region where the output adjustment processes are conducted are outputted in a specified format of file together with graphic images, etc., for example, in other photographic regions (S48). The output positions of character image, etc., including the layout style such as the left aligned, right aligned or centered, are specified by different manner depending on the file format.

**[0094]** Thus, in addition to the same effect as that of the first embodiment, the second embodiment provides a capability of placing the obtained character code data within the given character background region even when the character code data are outputted as a character image using a specific font so that it is possible to dispose the output character image within the same page without overlapping with other character images and/or graphic images while properly maintaining the original layout. Moreover, it is possible to layout character images maintaining the styling of each line, i.e., left alignment, right alignment and centering within the given character background regions.

**[0095]** It is obvious that this invention is not limited to the particular embodiments shown and described above but may be variously changed and modified without departing from the technical concept of this invention.

[0096] For example, although a digital copying machine having a capability of an image recognizing apparatus is used as an example in the above descriptions, the invention should not be construed to be limited to it. The invention can be applied to a computer that receives image data from a scanner and perform processes similar to the above-mentioned processes, a scanner that performs similar processes and transmit the results to a computer, printer, etc., a printer that receives image data directly from a scanner and performs similar processes, and other similar systems.

[0097] In the above embodiments, the image recognition controls is performed as the main control unit 11 executes a specific computer program in which the above processing sequence (see Fig. 2, Fig. 3, etc.) is described. This program can be offered in computer readable recording media such as flexible disks, CD-ROMs, etc. The program can also be provided as an application software program that executes the above process, or can be provided built into various information equipment such as a computer, a digital copier, and a control server.